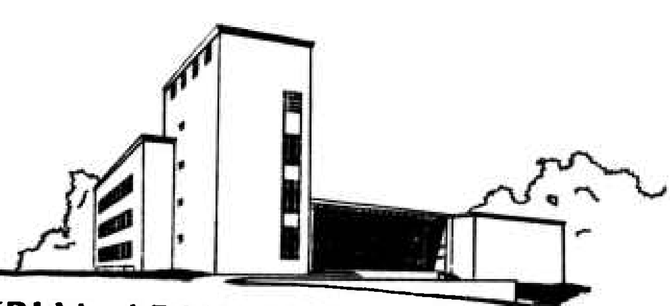
FC

#35

Carnegie Institute
of Technology

Pittsburgh 13, Pennsylvania



GRADUATE SCHOOL of INDUSTRIAL ADMINISTRATION

William Larimer Mellon, Founder

O. N. R. RESEARCH MEMORANDUM NO. 35

A NOTE ON THE NON-LINEAR POWER OF ADJACENT EXTREME POINT METHODS OF LINEAR PROGRAMMING*

by

A. Charnes, Purdue University and

W. W. Cooper, Carnegie Institute of Technology

September, 1955

Graduate School of Industrial Administration

Carnegie Institute of Technology

* This paper is based on earlier work done by the authors which was subsequently incorporated in the lecture notes of a course, "Set Theoretic Methods in Economic Analysis" offered by A. Charnes at the Graduate School of Industrial Administration, Carnegie Institute of Technology, during the spring semester, 1955.

Research was undertaken for the project Planning and Control of Industrial Operations, under contract with the Office of Naval Research. The paper is released for limited circulation in order to facilitate discussion, and invite criticism. It is tentative in character and should not be referred to in publications without the permission of the authors.

Background

1. Linear Programming and the Simplex Method

The general linear programming problem may be stated in the following form:

$$(1) \quad \text{maximize} \sum_{j=1}^{n} x_j \, c_j = z$$

$$\text{subject to} \sum_{j=1}^{n} a_{ij} \, x_j \leq b_i, \; i \; 1,2,\ldots, m.$$

$$x_j \geq 0, \; j \; 1,2,\ldots, n.$$

Research in the field of linear programming was originally directed to developing general methods for solving this problem. The development of the simplex method by Dantzig [7] provided a means for solving all such problems in a relatively efficient and straightforward manner. The process is as follows:

1. Inequalities are converted to equivalent equalities by adding suitably defined variables. A so-called tableau arrangement is then secured and an initial solution obtained which may be expressed as

   $$(2) \qquad \sum_{i=1}^{n} x_i \quad P_i = P_o \, ,$$

   where P_o represents the stipulations vector with components b_i, $i=1, 2, \ldots, m$, and the P_i are a suitably selected set of basis vectors.

2. The elements in the j^{th} column of the tableau may be regarded as coefficients y_{ij} for expressing P_j in terms of the P_i:

$$(3) \qquad \sum_{i=1}^{m} y_{ij} \; P_i = P_j \, , \qquad j=1, 2, \ldots, n.$$

3. An additional row is located with elements $z_j - c_j$ where

$$(4) \quad z_j = \sum_{i=1}^{m} y_{ij} \; c_i \, , \quad j=1,2,\ldots,n.$$

4. A criterion is supplied for judging the solutions at any stage. If all $z_j - c_j$ are non-negative, an optimum has been achieved; if not, further improvement is possible.

7

5. A systematic procedure is provided for changing the basis and continuing the calculations until an optimum is achieved. The procedure consists of designating a vector to come into the basis by reference to the $z_j - c_j$ and a vector to be removed by reference to the ratios x_i/y_{ij}, $y_{ij} > 0$.

2. Adjacent Extreme Point Methods

Since the simplex method was evolved numerous variants and alternative procedures (such as the dual method of Lemke [10]) have been devised. In common with the simplex method the procedures consist of movement from an extreme point to an adjacent extreme point of the convex set of solutions. The following properties of the solution set are therefore critical in all such procedures: (1) an optimum, if it exists, is attainable at an extreme point; (2) all optima can be generated from the extreme point optima. By proceeding along adjacent extreme points it is therefore possible to locate an optimum if one exists or to locate all optima, if desired. 1/

Extension to Non-Linear Functionals

Initially it was thought that these methods (which may be characterized as "adjacent extreme point methods") were applicable only to linear functionals with constraints given by linear inequalities. It was soon observed, however, that problems involving non-linear functionals of the form

$$(5) \qquad \sum_{k \in K} w_k \left| \sum_{r=1}^{n} a_{rk} \, x_r - S_k \right| , \qquad w_k \geq 0$$

which were to be minimized (subject to linear constraints) could be reduced to an equivalent linear programming problem. 2/ The reduction was achieved by introducing new variables

$$(6) \qquad x_k^{+}, \; x_k^{-} \geq 0$$

into the constraints and into the functional. Every optimum of the original (non-linear) problem could then be written as an optimum of the linear problem. Adjacent extreme point methods could therefore be used to locate the minimum values for this particular class of convex functionals.

Clearly, then, the power of adjacent extreme point methods extended to this class of convex functionals. It was thought, however, that problems involving minimization of more general classes of convex functionals -- e.g.,

---

1/ See Charnes [1]

2/ See [2]

minimization of semi-definite quadratic forms -- could not be solved by these methods except in special cases. Although, for such types of functionals the optimum is always on the boundary of the solution set, it is not necessarily located at an extreme point.

Subsequent research revealed that adjacent extreme point methods were capable of further extension. Charnes and Lemke [5] noted that problems involving minimization of separable convex functionals (and hence maximization of separable concave functionals) could be handled by these methods. Solutions (to any desired accuracy) could be achieved by means of piecewise linear approximations. Dantzig [6] showed further that this approach could be transformed into an equivalent formulation. A different extreme point method -- the bounded variables technique [4] and [8] -- could then also be used to solve such problems.

Still further extensions were soon forthcoming. In [9] Alan Hoffman announced that the methods of Charnes and Lemke for dealing with separable convex functionals had been extended to cover the general problem of minimizing an arbitrary convex functional subject to linear inequalities. He suggested, also, that this might be the limit to which the simplex method might be pushed in dealing with problems involving optimization of non-linear functionals subject to linear inequalities. Although full details of Dr. Hoffman's work are not yet available, his findings fit rather naturally into the evolution of research directed to exploring the boundaries of adjacent extreme point methods in solving linear programming problems.

Objectives

This paper is intended to carry these explorations a stage further. Specifically, the objectives are as follows: One, to show that it is possible to extend adjacent extreme point methods to a much wider class of non-linear functionals -- including functionals which need not be either convex or concave. Two, to characterize the class of such functionals which can be embraced by these methods. Three, to indicate modifications in the simplex procedures and criteria of optimality which are necessary to effect this extension. Four, to demarcate the limits of such extensions that may be made and thus establish the power of adjacent extreme point methods (including the simplex method) in dealing with optimization of arbitrary functionals subject to linear constraints.

Definition of Local Star Optima and Theorem

In order for extreme point methods to be effective (to the desired degree of approximation) it is clearly necessary for one of the following two conditions to hold: (1) the optima are at extreme points or (2) the problem must be transformable into an equivalent one in which the corresponding optima of the original problem are at extreme points of the new problem. The behavior of the functional at points other than extreme points is irrelevant.

Since the issue is optimization in the large these conditions, though necessary, are not sufficient. Adjacent extreme point methods may, when these conditions are fulfilled, yield only local optima. It will be convenient therefore to define

> A Local Star Optimum: An extreme point solution which yields a functional value at least as great as can be attained at any adjacent extreme point.

Clearly then the following theorem is valid:

> Theorem: A necessary and sufficient condition that adjacent extreme point methods always yield an optimum is that all local star optima shall also be optima in the large.

Illustration

The following illustration will serve to show what is involved and provide a basis for subsequent discussion. Suppose that the objective is to minimize the non-linear functional

$$(7) \qquad F\ (y_1, y_2, \ldots, y_n)$$

subject to a set of linear constraints. It is possible to approximate F to any desired degree of accuracy by means of piecewise linear functionals, provided F may be decomposed by linear transformations into a sum of functions of individual variables. Thus, if F can be written as

$$(8) \qquad \sum_{i=1}^{p} f_i\ (x_i)$$

where

$$x_i = \sum_{k=1}^{n} a_{ik}\ y_k$$

then the f_i (and hence F) can be approximated (to any desired degree) by piecewise linear functionals.

Consider, for example, the function $f(x)$ sketched in Figure I

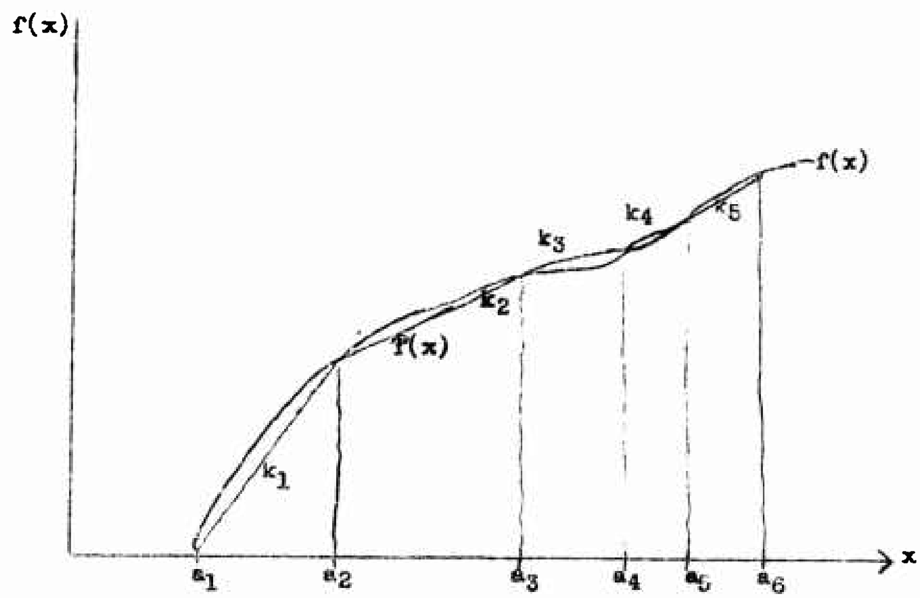


FIGURE I

PIECEWISE LINEAR APPROXIMATION
TO CONCAVE AND CONVEX FUNCTION

which is convex over certain domains and concave over others.[3/] The piecewise linear approximation is given by $\hat{f}(x)$. The k's represent the slopes of the approximating lines and the a's represent the initial and terminal points of each successive segment.

3/ Continuity is not essential.

Numerous convenient devices are available for securing the desired degree of approximation. The following procedure provides an example when f is differentiable. The value $\frac{df}{dx}$ is plotted against x, as in Figure II. The area under $\frac{df}{dx}$ may be used to represent $f(x)$ and the histogram used to represent $\hat{f}(x)$. By choosing appropriate a_i's $(i=1,2,\ldots,n)$ the absolute difference $\left|\frac{f(x)-\hat{f}(x)}{f(x)}\right|$ may be made as small as desired and a suitable approximation thus obtained.

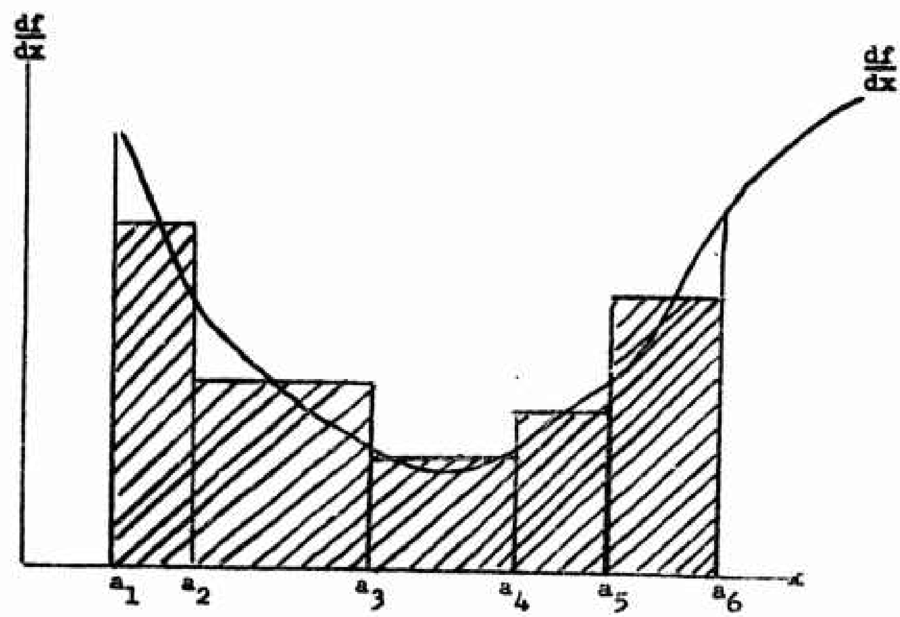


FIGURE II

HISTOGRAM APPROXIMATION TO $f(x)$

Analytically this approximation may be written as follows.

$$
\text{(9)} \quad \hat{f}(x) = \begin{cases}
0 & , \ x \le a_1 \\
k_1\ (x-a_1) & , \ a_1 \le x \le a_2 \\
k_1\ (x-a_1) + (k_2 - k_1)(x-a_2) & , \ a_2 \le x \le a_3 \\
k_1\ (x-a_1) + (k_2 - k_1)(x-a_2) + (k_3-k_2)(x-a_3), & a_3 \le x \le a_4 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
k_1\ (x-a_1) + (k_2 - k_1)(x-a_2) + \ldots + (k_6-k_5)(x-a_6), & a_6 \le x.
\end{cases}
$$

An alternative expression is

$$
\text{(10)} \quad \hat{f}(x) = \sum_{j=1}^{m} (k_j - k_{j-1})\ (x - a_j), \quad a_m \le x \le a_{m+1}
$$

where $k_o = 0$.

Setting

$$x_i^+ - x_i^- = x - a_i$$

where

$$
\text{(11)} \quad \begin{aligned} & x_i^+ \ , \ x_i^- \ge 0 \\ & x_i^+ \quad x_i^- = 0, \end{aligned}
$$

$i = 1, 2, \ldots, 6$

the expression (10) may be written as,

$$
\text{(12)} \quad \begin{aligned} f(\hat{x}) &= \sum_{j=1}^{6} (k_j - k_{j-1})\ (x_j^+ - x_j^-) \ , \ a_m \le x \le a_{m+1} \\ &= \sum_{j=1}^{6} (k_j - k_{j-1})\ x_j^+ \end{aligned}
$$

Modifications of Simplex Criteria and Procedures

As the example of the preceding section illustrates the functional F may be decomposed by linear transformations into a sum of functionals each involving only a single variable. To each such functional a corresponding (approximating) linear functional may be obtained which is to be optimized in an enlarged domain involving additional linear equations and specified quadratic conditions in non-negative variables.

The linear functionals are of the form

$$(13) \qquad \hat{f}_j(x_j) = \sum_{s=1}^{n} (k_{js} - k_{j(s-1)})\, x^+_{js}.$$

$s=1,2,\ldots,n'$.

The additional linear equations which must then be adjoined to the original set of restrictions is

$$(14) \qquad \begin{array}{lccl} x_j - (x^+_{j1} - x^-_{j1}) & & & = a_{j1} \\ x_j & - (x^+_{j2} - x^-_{j2}) & & = a_{j2} \\ \vdots & & \ddots & \vdots \\ x_j & & -(x^+_{jn} - x^-_{jn}) & = a_{jn} \end{array}$$

Finally, the following quadratic conditions must be satisfied

$$(15) \qquad \begin{array}{l} x^+_{j1}\ x^-_{j1} = 0 \\ x^+_{j2}\ x^-_{j2} = 0 \\ \quad\ldots \\ x^+_{jn}\ x^-_{jn} = 0 \end{array}$$

The coordinates at extreme points, however, are coefficients of linearly independent sets of vectors.[4] But as the expressions in (14) clearly show the vector coefficients of x_{js}^{+} and x_{js}^{-} are linearly dependent. Hence, as long as movement is along extreme points only, either $x_{js}^{+} = 0$ or $x_{js}^{-} = 0$ so that the quadratic conditions (15) are automatically satisfied and the representation (13) of the functional is valid. It is therefore necessary to stipulate that the method of calculation utilizes extreme points only.

If maximization is being undertaken and if the functional is convex over certain ranges the ordinary $z_j - c_j$ improvement criteria must be modified because of the possible appearance of infinite values for the functional.[5] To fix ideas consider a problem involving a non-linear functional which has been reduced to the form

$$\max. \sum_{j=1}^{n+n'} c_j \lambda_j$$

(16) subject to

(a) $$\sum_{j=1}^{n} a_{ij} \lambda_j = b_i , \quad i\ 1,2,\ldots,m$$

(b) $$\sum_{j=1}^{n+n'} \gamma_{(i+m)j} \lambda_j = \gamma_{i+m}, \quad i=1,2,\ldots,m'$$

(c) $$\lambda_j \geq 0$$

4/ See [3]

5/ See rule 4, section 1. Similar remarks apply if minimization is being undertaken and the functional is concave over certain ranges.

where the restrictions (a) are associated with the original problem and the restrictions (b) are adjoined (as in the illustration of the preceding section) in the process of securing the approximation. Note: (1) that the object in the approximation is to obtain a value which agrees with $f_j(x_j)$ on the extreme points, (2) that different numbers of subdivisions may be used for each $f_j(x_j)$, and (3) so long as movement is restricted to extreme points the quadratic condition is automatically satisfied and may thus be regarded as redundant.

In more compact vector notation, (16) may be reexpressed as

$$\text{(17)} \qquad \max. \sum_{j=1}^{n+n'} c_j \lambda_j$$

subject to

$$\sum_{j=1}^{n} \begin{pmatrix} P_j \\ D_j \end{pmatrix} \lambda_j + \sum_{j=n+1}^{n'} \begin{pmatrix} 0 \\ G_j \end{pmatrix} \lambda_j = \begin{pmatrix} P_o \\ G_o \end{pmatrix}$$

$$\lambda_j \geq 0$$

where the P's, D's and G's, (as well as 0) are appropriately defined column vectors of m and m' rows respectively.

Evidently any linearly independent set of m P_j's must be associated with a set of m' linearly independent G_j's if movement is to be restricted to extreme points. Hence rule 5 of the simplex routine which stipulates replacement of a P_j by a P_k to provide a succeeding basis must be modified so that a $\begin{pmatrix} P_j \\ D_j \end{pmatrix}$ can only be replaced by a $\begin{pmatrix} P_k \\ D_k \end{pmatrix}$ and $\begin{pmatrix} 0 \\ G_j \end{pmatrix}$ by a $\begin{pmatrix} 0 \\ G_k \end{pmatrix}$. The optimality criterion must then be modified to: $z_j - c_j \geq 0$ for all permissible replacements.

By applying these rules to the extended set, movement is restricted to extreme points of the original set. Hence, an optimum which is attained for the enlarged problem of employing these altered procedures and criteria is also an optimum for the original problem. In general, only a local star optimum is guaranteed. When a local star optimum is also an optimum in the large the procedures and criteria outlined above will suffice to locate this point irrespective of the non-linear characteristics of the functional.

Further Extensions

The procedure is capable of further extension in particular applications. Clearly all cases in which local star optima are optima in the large are comprehended by the above rules and criteria. In other cases it is possible to employ these new criteria to reach various local star optima by starting at widely separated extreme points. It may then be possible to establish that an optimum optimorum has been attained in the process.

The ability to use adjacent extreme point methods (such as the simplex procedure) as a general procedure for solving all such problems is patently dependent on the development of systematic and efficient means for traversing local star optima. At present no such general techniques are available. The procedure outlined above is, however, amenable to heuristic employment when a priori considerations do not make it apparent that a local star optimum is also an optimum in the large. On attaining a local star extremum it may be possible to establish directly that it is also an optimum optimorum. The current lack of general and efficient traversal techniques demarcates, for the present, the non-linear power of adjacent extreme point methods of linear programming.

BIBLIOGRAPHY

[1] Charnes, A. "Degeneracy and Optimality in Linear Programming," *Econometrica* XX, 2.

[2] ———, W. W. Cooper and R. Ferguson, "Optimal Estimation of Executive Compensation by Linear Programming," *Management Science* Vol. 1, No. 2

[3] Charnes, A., W. W. Cooper and A. Henderson, *An Introduction to Linear Programming* (New York: John Wiley and Sons, Inc., 1953)

[4] Charnes, A. and C. E. Lemke, "Computational Theory of Linear Programming I: The 'Bounded Variables' Problem," ONR Research Memorandum No. 10, (Pittsburgh: Graduate School of Industrial Administration, Carnegie Institute of Technology, Jan., 1954)

[5] ———, "Minimization of Non-Linear Separable Convex Functionals," *Naval Research Logistics Quarterly*, Vol. 1, No. 4.

[6] Dantzig, G. B., "Developments in Linear Programming," *Management Science* (forthcoming)

[7] Dantzig, G. B., "Maximization of a Linear Function of Variables Subject to Linear Inequalities," *Activity Analysis of Production and Allocation*, T. C. Koopmans, editor.

[8] Dantzig, G. B., "Upper Bounds, Secondary Constraints and Block Triangularity in Linear Programming," *Econometrica* Vol. 23, No. 2

[9] Hoffman, A. J., Quarterly Report, July through September, 1954 (Washington: National Bureau of Standards)

[10] Lemke, C. E., "The Dual Method of Solving the Linear Programming Problem," *Naval Research Logistics Quarterly*, Vol. 1, No. 1, March, 1954